Object extraction using a stochastic birth-and-death dynamics in continuum

Xavier Descombes, Robert Minlos, Elena Zhizhina

Object extraction using a stochastic birth-and-death dynamics in continuum

Xavier Descombes Robert Minlos Elena Zhizhina *†

September 16, 2008

[1]

Abstract

We define a new birth and death dynamics dealing with configurations of disks in the plane. We prove the convergence of the continuous process and propose a discrete scheme converging to the continuous case. This framework is developed to address image processing problems consisting in detecting a configuration of objects from a digital image. The derived algorithm is applied for tree crown extraction and bird detection from aerial images. The performance of this approach is shown on real data.

Keywords: Object extraction, Stochastic modeling, Birth and death dynamics

*X. Descombes is within Ariana Research Group, INRIA/I3S, 2004 route des Lucioles, BP 93, 06902, Sophia Antipolis cedex, France

†R. Minlos and E. Zhizhina are within Lab. Dobrushin- IITP, Bol'shoi Karetnyi per., 19, 127994, GSP-4, Moscow, Russia.

[1]This work was partially supported by EGIDE within the ECO-NET project 18902PK, by INRIA COLORS "Flamants" and by the INRIA Associated team "ODESSA". We would like to thank the French National Forest Inventory (IFN) and Arnaud Béchet from "La Station Biologique Tour du Valat" for kindly providing the data. R. Minlos and E. Zhizhina are partially supported by RFFI grant 08-01-00105a.

1 Introduction

Object detection from remotely sensed images, such as aerial or satellite images, is a crucial step in many applications, including change detection or counting a targeted population. When addressing high resolution images, the pixel itself does not embed all the information. In this context, taking into account the geometrical information is challenging. In this paper, we therefore consider a model based on objects rather than on pixels. Embedded in a stochastic framework, such models increases the complexity of optimization algorithms. The speed of convergence of the underlying dynamics is a major issue for the relevance of the approach in case of concrete applications. In this context, we propose a new stochastic algorithm to solve object extraction problems from images which outperforms the convergence speed of RJMCMC algorithms. In this paper, we consider here a model of possibly partially overlapping disks. Each disk in the final configuration is associated with a given object in the image, for example a tree or a bird. We define a continuous time stochastic evolution of a set of objects, in the image plan converging to the set of objects of interest in the image.

The evolution under consideration is a birth-and-death equilibrium dynamics on the configuration space of disks (or the configuration space of points) with a given stationary Gibbs measure (see [1, 2]). We define birth and death rates according to the so-called detailed balance conditions. In our scheme the intensity of birth is constant, whereas intensities of death depend on both the energy function and the current configuration. This choice of rates has been made to optimize the convergence speed. Indeed, the volume of the space for birth, equal to the number of pixels in the image under study, is much bigger than the number of disks in the configuration. It is therefore faster to update the death map than the birth map.

We then embed the defined stationary dynamics into a simulated annealing procedure where the temperature of the system tends to zero in time. We thus obtain a non-stationary stochastic process, such that all weak limit measures have a support on configurations giving the global minimum of the energy function under a minimal number of disks in the configuration. The final step is the discretization of this non-stationary dynamics. The resulting discrete process is a non-homogeneous (in time and in space) Markov chain with transition probabilities depending on the temperature, the energy function and the discretization step. Hereafter we prove that:

1) the discretization process converges to the continuous time process under fixed temperature as the step of

discretization tends to zero;

2) if we apply the discretization process to any initial measure with a continuous density w.r.t. the Lebesgue-Poisson measure, then in the limit when the discretization step tends to 0, time tends to infinity and the temperature tends to 0, we get a measure concentrated on the global minima of the energy function with a minimal number of disks.

In section 5, we apply this framework to object detection from images. Marked point processes, defined by a Gibbs measure against the Poisson process, have proved to be adapted to such problems by modeling simple geometric objects defined by the marks associated to each point. Moreover, interactions between points allow to model some a priori information on the object configuration. This approach have been applied to detect different features such as road networks [3, 4], buildings [5] or trees [6]. In these references, the optimization of the defined marked point process is performed using a RJMCMC scheme [7]. In the RJMCMC scheme, each iteration consists in perturbing one or a couple of objects. Besides the rejection rate induces a huge computation time. The main point of our approach is that each step concerns the whole configuration and there is no rejection. We thus obtained better performances in term of computational time, which allow to deal with real images of several millions of pixels. We build an energy function which embeds some a priori knowledge on the object configuration such as partial non overlapping between objects and a data term which allows the objects to fit the image under study. The optimization is then performed by using the proposed birth and death dynamics. Some results are shown on real data for the two different problems concerning tree and bird detection.

2 Description of the model

In this section, we give the mathematical background underlying the proposed algorithm by introducing the configuration space, the Gibbs distribution on this configuration space, and all necessary constructions for the continuous time dynamics presented in the next section.

2.1 Configuration space

We consider finite sets of disks $\{d_{x_1}, \ldots, d_{x_k}\}$ of the same radius r, lying in a bounded domain $V \subset R^2$, with a hard core distance ϵ, i.e. the minimum distance between any pair of disk center is ϵ, . Let

$$\gamma = \{x_i\} \in \Gamma_d(V), \quad x_i \in V \subset R^2,$$

be a configuration defined by a non-ordered set of points, each point representing the center of a disk. $\Gamma_d(V)$ denotes the configuration space given by the set of finite sets of points in V. The space $\Gamma_d(V)$ can be endowed with the weakest topology with respect to which all maps $\gamma \to < f, \gamma > = \sum_{x \in \gamma} f(x), f \in C_0^\infty(R^2)$ are continuous, $C_0^\infty(R^2)$ being the space of infinitely differentiable real-valued functions on R^2. Since the domain V is bounded, the number of disks (or equivalently the number of points) in any configuration is uniformly bounded

$$|\gamma| \ < \ N \ = \ \frac{4|V|}{\pi\epsilon^2},$$

where $|V|$ is the volume of V. Let denote $\Gamma_d(V, n)$ the set of configurations containing exactly n points. Then $\Gamma_d(V)$ is defined as the union of the $\Gamma_d(V, n)$:

$$\Gamma_d(V) \ = \ \bigcup_{n=0}^{N} \Gamma_d(V, n),$$

$\Gamma_d(V, 0) = \{\emptyset\}$ corresponds to the configuration with no disk. Since we consider unordered sets of disks, the set $\Gamma_d(V, n)$, for any $n > 0$, can be represented as a quotient space

$$\Gamma_d(V, n) \ = \ V_d^n \ / \ S_n,$$

where

$$V_d^n = \{(x_1, \ldots, x_n) \in V^n : \ |x_i - x_j| \geq \epsilon, \ i, j = 1, \ldots, n, \ i \neq j\},$$

and S_n is the permutation group in the set $(x_1, \ldots, x_n)$. We define a mapping

$$\Pi_n : V_d^n \ \to \ \Gamma_d(V, n), \quad \Pi_n(x_1, \ldots, x_n) \in \Gamma_d(V, n) \tag{1}$$

Since the number of points in configurations is random, it is natural to consider the functional space in the form of a Fock space, where different components of a function may depend on different number of variables. Therefore, each function $F(\gamma), \ \gamma \in \Gamma_d(V)$ on the space $\Gamma_d(V)$ can be represented as follows:

$$F_0, \ F_1(x_1), \ F_2(x_1, x_2), \ldots, F_N(x_1, \ldots, x_N), \tag{2}$$

where $F_0 = F(\emptyset),\ F_n(x_1,\dots,x_n) = F(\Pi_n(x_1,\dots,x_n))$ is a symmetrical function on V_d^n. A function F on the space $\Gamma_d(V)$ is said to be a smooth (continuous) function if each function in (2) is a smooth (continuous) function on V_d^n.

Let us consider a measure in the space $\Gamma_d(V,n)$:

$$\lambda_n(A) \ = \ \frac{\left|\Pi_n^{-1}(A)\right|}{n!}, \quad A \subset \Gamma_d(V,n),\ \lambda_0(\emptyset) = 1. \tag{3}$$

Here $|\Pi_n^{-1}(A)|$ is the $2n$-dimensional Lebesgue volume of the domain $\Pi_n^{-1}(A) \subset V_d^n$ (a complete preimage of A under the mapping Π_n). The measure λ on the space $\Gamma_d(V)$, such that the restriction of λ on each set $\Gamma_d(V,n)$ is given by λ_n, is called the Lebesgue-Poisson measure.

2.2 Energy function

We define on the space $\Gamma_d(V)$ a real-valued smooth and bounded from below function $H(\gamma)$ which is called the energy function. We set $H(\emptyset) = 0$. Then the corresponding Fock representation for H has the form

$$H \ = \ (0,\ H_1(x_1),\ H_2(x_1,x_2),\dots,H_N(x_1,\dots,x_N))\,. \tag{4}$$

The Gibbs distribution μ_β^V on the space $\Gamma_d(V)$ generated by the energy $H(\gamma)$ is defined by the density $p_V(\gamma) = \frac{d\mu_\beta^V}{d\lambda}(\gamma)$ with respect to the Lebesgue-Poisson measure λ:

$$p_V(\gamma) \ = \ \frac{z^{|\gamma|}}{Z_{\beta,V}}\ \exp\{-\beta H(\gamma)\}, \tag{5}$$

with positive parameters $\beta > 0,\ z > 0$ and a normalizing factor $Z_{\beta,V}$:

$$Z_{\beta,V} = \int\limits_{\Gamma_d(V)} z^{|\gamma|} \exp\{-\beta H(\gamma)\} d\lambda(\gamma) = 1 + \sum_{n=1}^{N} \frac{z^n}{n!} \int\limits_{V_d^n} e^{-\beta H_n(x_1,\dots,x_n)} dx_1 \dots dx_n.$$

We formulate now some assumptions on the energy function $H_V(\gamma)$. Denote by

$$\bar{H} \ = \ \min_{\gamma \in \overline{\Gamma_d(V)}} H(\gamma),$$

where $\overline{\Gamma_d(V)}$ is the closure of $\Gamma_d(V)$, and let

$$T_V \ = \ \{\bar{\gamma} \in \overline{\Gamma_d(V)} :\ H(\bar{\gamma}) = \bar{H}\}$$

be the set of all points in $\overline{\Gamma_d(V)}$ giving the global minimum $\bar{H}$ of the function $H(\gamma)$. The set T_V can be written as

$$T_V \ = \ \bigcup_{n=0}^{N} T_{V,n},$$

where $T_{V,n}$ is a set of configurations from T_V which are also configurations from $\Gamma_d(V,n)$, i.e. contain exactly n disks.

In practice, this energy contains a first term representing a priori knowledge on the disks configuration and which is defined by interactions between neighboring disks, and a second term, obtained from data, which is defined for each object and which can be negative (see e.g. equation (32)).

2.3 Convergence of measures

We first establish a result necessary to define an annealing scheme for optimizing the proposed model, i.e. for finding a configuration minimizing the energy function. We assume that

1) the set T_V is finite and situated in $\Gamma_d(V)$,

2) for any configuration $\bar{\gamma} \in T_{V,n}$ any preimage $(\bar{x}_1, \ldots, \bar{x}_n) \in \Pi_n^{-1}(\bar{\gamma})$ of $\bar{\gamma}$ is a non-degenerated critical point of the function H_n, which depends on n points or equivalently $2n$ variables, i.e.:

$$\frac{\partial H_n}{\partial x_i^{(m)}}(\bar{x}_1, \ldots, \bar{x}_n) \ = \ 0, m = 1, 2$$

for any $i = 1, \ldots, n$, $\bar{x}_i = (x_i^{(1)}, x_i^{(2)}) \in V$ are coordinates of point $\bar{x}_i$, and the matrix $\mathcal{A}$ of size $2n \times 2n$ defined as follows:

$$\mathcal{A}\ (\bar{\gamma}) \ = \ \left\{ \frac{\partial^2 H_n}{\partial x_i^{(m_i)} \partial x_j^{(m_j)}}(\bar{x}_1, \ldots, \bar{x}_n) \right\}, m_i, m_j = 1, 2,$$

at point $(\bar{x}_1, \ldots, \bar{x}_n)$ is strictly positive-definite (this matrix is the same for all preimages of the configuration $\bar{\gamma}$). We denote

$$B(\bar{\gamma}) \ = \ det\ \mathcal{A}(\bar{\gamma}). \tag{6}$$

Theorem 1. *Let $n_0 \in [0, \dots, N]$ be the minimal index for which the set $T_{V,n}$ is not empty. Then the Gibbs distributions μ_β converge weakly as $\beta \to \infty$ to a distribution μ_∞ on $\Gamma_d(V)$ of the form*

$$\mu_\infty \;=\; \sum_{\bar\gamma \in T_{V,n_0}} C_{\bar\gamma}\delta_{\bar\gamma} \text{ if } n_0 > 0, \quad \text{and} \quad \mu_\infty \;=\; \delta_{\{\emptyset\}} \text{ if } n_0 = 0. \tag{7}$$

Here $\delta_{\bar\gamma}$ is the unit measure concentrated on the configuration $\bar\gamma$, and the coefficients $C_{\bar\gamma}$ satisfy the following equality

$$\sum_{\bar\gamma \in T_{V,n_0}} C_{\bar\gamma} \;=\; 1.$$

Proof. Let $F(\gamma)$ be a smooth function on $\Gamma_d(V)$. Then

$$\langle F\rangle_{\mu_\beta} \;= \frac{I(F)}{Z_{\beta,V}} = Z^{-1}_{\beta,V}\left(F_0 \;+\; \sum_{n=1}^{N}\frac{z^n}{n!}\int\limits_{V_d^n} F_n(x_1,\dots,x_n)\; e^{-\beta H_n(x_1,\dots,x_n)}\; dx_1\dots dx_2\right). \tag{8}$$

Let us consider each integral in (8)

$$I_n(F_n) \;=\; \int\limits_{V_d^n} F_n(x_1,\dots,x_n)\; e^{-\beta H_n(x_1,\dots,x_n)}\; dx_1\dots dx_2$$

separately. If the set $T_{V,n}$ is empty, then the integral $I_n(F_n)$ is bounded as follows:

$$|I_n(F_n)| \;<\; e^{-\beta h_n}|V_d^n|, \tag{9}$$

where

$$h_n \;=\; \min H_{V,n}(x_1,\dots,x_n) \;>\; \bar H.$$

If $T_{V,n}$ is not empty, then the following asymptotic holds as $\beta \to \infty$, see for example [9] or [8], chapter II.

$$I_n(F_n) \;=\; e^{-\beta\bar H}\left(\sum_{\bar\gamma\in T_{V,n}} F(\bar\gamma)R(\bar\gamma)\beta^{-n/2} \;+\; \beta^{-n/2-1/2}S(F,\beta)\right), \tag{10}$$

where $S(F,\beta)$ is bounded as $\beta\to\infty$, and

$$R(\bar\gamma) \;=\; \frac{1}{(2\pi)^{n/2}}B^{-1/2}(\bar\gamma),$$

where $B(\bar\gamma)$ is defined in (6). Then bound (9) and asymptotic (10) imply that for $n_0 > 0$ and $\beta\to\infty$ we have

$$I(F) \;=\; F_0 \;+\; \sum_{n=1}^{N}\frac{z^n}{n!}I_n(F_n) \;=\; \frac{e^{-\beta\bar H}}{\beta^{n_0/2}}\left(\sum_{\bar\gamma\in T_{V,n_0}} F(\bar\gamma)R(\bar\gamma) \;+\; o(1)\right). \tag{11}$$

Analogously,

$$Z_{\beta,V} \;=\; I(1) \;=\; \frac{e^{-\beta\bar H}}{\beta^{n_0/2}}\left(\sum_{\bar\gamma\in T_{V,n_0}} R(\bar\gamma) \;+\; o(1)\right). \tag{12}$$

Finally from (8), (11) and (12) we get (7) with

$$C(\bar{\gamma}) \;=\; \frac{R(\bar{\gamma})}{\sum_{\bar{\gamma}\in T_{V,n_0}} R(\bar{\gamma})}.$$

The case $n_0 = 0$ can be studied in the same way. Since the space of smooth functions is dense in the space of bounded functions $B(\Gamma_d(V))$, we have proved the weak convergence of measures $\mu_\beta \to \mu_\infty$ on the space $B(\Gamma_d(V))$. □

3 A continuous-time equilibrium dynamics

We now define a dynamics to simulate the proposed model. Generally, the natural description of a continuous time dynamics can be given by its generator, see e.g. [11]. In our case, the generator of a birth-and-death process in the domain $V \subset R^2$ is an operator in the space $B(\Gamma_d(V))$ of bounded measurable functions on $\Gamma_d(V)$ of the form (see also [2] for details):

$$(L_\beta\ f)(\gamma) \;=\; \sum_{x\in\gamma} e^{\beta E(x,\gamma\backslash x)}(f(\gamma\backslash x) - f(\gamma)) \;+\; z \int\limits_{V(\gamma)} (f(\gamma\cup y) - f(\gamma))\ dy, \tag{13}$$

where

$$V(\gamma) \;=\; V\backslash D(\gamma), \quad D(\gamma) \;=\; (\cup_{x\in\gamma}\mathcal{B}_x(\epsilon))\cap V,$$

$\mathcal{B}_x(\epsilon)$ being the disk with center at point $x \in V$ and radius ϵ, and

$$E(x,\gamma\backslash x) \;=\; H(\gamma) - H(\gamma\backslash x). \tag{14}$$

The birth intensity $b(\gamma, x)$, consisting in adding the point x to the configuration γ is given by:

$$b(\gamma,x)\ dx \;=\; z\ dx \text{ if } x \notin D(\gamma). \tag{15}$$

The death intensity $d(\gamma, x)$ which consists in removing the point x from the configuration γ is given by:

$$d(\gamma,x) \;=\; e^{\beta E(x,\gamma\backslash x)} \text{ if } x \in \gamma. \tag{16}$$

Under this choice for the birth and death intensities, the detailed balance condition holds:

$$\frac{b(\gamma,x)}{d(\gamma,x)} \;=\; \frac{p_V(\gamma)}{p_V(\gamma\backslash x)} \;=\; z\ e^{-\beta E(x,\gamma\backslash x)},$$

and consequently, see for example [10], the corresponding birth-and-death process associated with the stochastic semi-group $T_\beta(t) = e^{tL_\beta}$ is time reversible, and its equilibrium distribution is the Gibbs stationary measure μ_β^V with density given by equation (5).

Theorem 2. 1) *The operator L_β is a bounded operator in the space of bounded functions $B(\Gamma_d(V))$ and in $L_2(\Gamma_d(V), \mu_\beta)$, moreover L_β is a self-adjoint bounded operator in $L_2(\Gamma_d(V), \mu_\beta)$.*

2) *The family of operators $T_\beta(t) = e^{tL_\beta},\ t \geq 0$ forms a self-adjoint Markov semi-group in $L_2(\Gamma_d(V), \mu_\beta)$, i.e.*

$$e^{tL_\beta} 1 = 1, \quad \text{and} \quad e^{tL_\beta} F \geq 0 \ \text{ for any non-negative } \ F \in L_2(\Gamma_d(V), \mu_\beta). \tag{17}$$

3) *The semi-group $T_\beta(t)$ is a contraction semi-group in $B(\Gamma_d(V))$.*

For the proof, see Appendix A.

Corollary. Property 2 implies that there exists a unique fixed vector of the operators $T_\beta(t) = e^{tL_\beta}$ in $L_2(\Gamma_d(V), \mu_\beta)$ (which is equal to 1), see [12].

The convergence to the stationary measure μ_β is guaranteed by the general result given by C. Preston in [13]. We consider a family $\mathbb{B}(\lambda)$ of measures ν on the space $\Gamma_d(V)$ with a bounded density $\tilde{p}_\nu(\gamma)$ with respect to the Lebesgue-Poisson measure λ. This, in particular, implies that the density $p_\nu(\gamma)$ of the measure $\nu \in \mathbb{B}(\lambda)$ w.r.t. a Gibbs measure μ_β (for any β):

$$p_\nu(\gamma) \ = \ \frac{d\nu}{d\mu_\beta}(\gamma)$$

is also bounded, and consequently, $p_\nu(\gamma) \in L_2(\Gamma_d(V), \mu_\beta)$. Then we can define the evolution $\nu_t \equiv T(t)\nu$ of the measure $\nu \in \mathbb{B}(\lambda)$ as follows:

$$\langle \nu_t, F \rangle \ = \ (T_\beta(t) p_\nu, F)_{\mu_\beta}.$$

Notice that property 2) of Theorem 2 implies that $T_\beta(t)p_\nu$ is again a density w.r.t. a Gibbs measure, i.e.

$$T_\beta(t)p_\nu \geq 0, \quad \langle T_\beta(t)p_\nu \rangle_{\mu_\beta} = \langle p_\nu \rangle_{\mu_\beta} = 1.$$

Theorem 3. *Let $\nu \in \mathbb{B}(\lambda)$. Then for any $F \in L_2(\Gamma_d(V), \mu_\beta)$ we get*

$$\langle T_\beta(t)\nu, F \rangle \ \equiv \ \langle \nu_t, F \rangle \ = \ (T_\beta(t) p_\nu, F)_{\mu_\beta} \ \to \ \langle F \rangle_{\mu_\beta}. \tag{18}$$

The proof of theorem 3 follows from the general theorems by C. Preston [13].

4 Approximation process

In practice, for computer simulations, we need to discretize the process in time. In this section, we construct a discrete time approximation of the proposed continuous birth and death process.

We consider Markov chains $T_{\beta,\delta}(m), m = 0, 1, 2, \ldots$ on the same space $\Gamma_d(V)$. The process $T_{\beta,\delta}(m)$ is defined as follows: a configuration γ is transformed into a configuration $\gamma' = \gamma_1 \cup \gamma_2$, where $\gamma_1 \subseteq \gamma$, and γ_2 is a configuration of points such that $\gamma_1 \cap \gamma_2 = \emptyset$ and is distributed w.r.t. the Poisson law with intensity z.

This transformation embed a birth part given by γ_2 and a death part given by $\gamma \backslash \gamma_1$.

The transition probability for the death of a point x (i.e. a disk with the center at x) from the configuration γ is given by:

$$p_{x,\delta} = \begin{cases} \frac{e^{\beta E(x,\gamma\backslash x)}\,\delta}{1 + e^{\beta E(x,\gamma\backslash x)}\,\delta} = \frac{a_x\delta}{1+a_x\delta}, & \text{if} \quad \gamma \to \gamma\backslash x, \\ \\ \frac{1}{1+a_x\delta}, & \text{if} \quad \gamma \to \gamma \quad (x \text{ survives}). \end{cases} \tag{19}$$

where $a_x = a_x(\gamma) = e^{\beta E(x,\gamma\backslash x)}$, $E(x, \gamma\backslash x)$ being defined by equation (14). Moreover, all the points are killed independently, and the configurations γ_1 and γ_2 are independent.

The transitions associated with the birth of a new point in a small domain $\Delta y \subset V(\gamma)$ have the following probability distribution:

$$q_{y,\delta} = \begin{cases} z\Delta y\delta, & \text{if} \quad \gamma \to \gamma \cup y, \\ \\ 1 - z\Delta y\delta, & \text{if} \quad \gamma \to \gamma \quad (\text{ no birth in } \Delta y). \end{cases} \tag{20}$$

Using these definitions, the transition operator $P_{\beta,\delta}$ for the process, $T_{\beta,\delta}(m) = P^m_{\beta,\delta}$, has the following form:

$$\begin{aligned} \left(P_{\beta,\delta} f\right)(\gamma) &= \sum_{\gamma_1 \subseteq \gamma} \prod_{x \in \gamma_1} \frac{1}{1 + a_x\delta} \prod_{x \in \gamma\backslash\gamma_1} \frac{a_x\delta}{1 + a_x\delta} \\ \Xi^{-1}_{\delta}(\gamma_1) &\sum_{k=0}^{\infty} \int\limits_{V_k(\gamma_1)} \frac{(z\delta)^k}{k!} f(\gamma_1 \cup y_1 \cup \ldots \cup y_k)\, dy_1 \ldots dy_k, \end{aligned} \tag{21}$$

where $\Xi_\delta(\gamma) = \Xi_\delta(V(\gamma), z, \delta)$ is the normalizing factor for the conditional Lebesgue-Poisson measure under a given configuration of points γ_1. We prove below that the approximation process $T_{\beta,\delta}(m) \equiv T_{\beta,\delta}\left(\left[\frac{t}{\delta}\right]\right)$ converges to the continuous time process $T_\beta(t)$ uniformly on bounded intervals $[0, \bar{t}]$ as the discretization step δ tends to 0.

Let us denote $\mathfrak{L} = B(\Gamma_d(V))$ the Banach space of bounded functions on Γ_V with the norm

$$\|F\| = \sup_{\gamma \in \Gamma_d(V)} |F(\gamma)|.$$

Theorem 4. *For each* $F \in \mathfrak{L}$

$$\|T_{\beta,\delta}(t)F - T_\beta(t)F\|_{\mathfrak{L}} = \sup_\gamma |(T_{\beta,\delta}(t)F)(\gamma) - (T_\beta(t)F)(\gamma)| \to 0, \tag{22}$$

as $\delta \to 0$ *for all* $t \geq 0$ *uniformly on bounded intervals of time.*

See Appendix B for the proof.

Corollary. The result of theorem 4 implies that for any $F, G \in B(\Gamma_d(V))$ we have

$$(G, T_{\beta,\delta}(t)F)_{\mu_\beta} \to (G, T_\beta(t)F)_{\mu_\beta} \quad \text{as} \quad \delta \to 0. \tag{23}$$

We denote by $S_{\beta,\delta}(m)$ an adjoint to $T_{\beta,\delta}(m)$ semi-group acting on measures, such that for any $\nu \in \mathbb{B}(\lambda)$:

$$\langle S_{\beta,\delta}(m)\nu, F\rangle = (p_\nu, T_{\beta,\delta}(m)F)_{\mu_\beta}.$$

We now formulate the main result about convergence, when both parameters δ and $\frac{1}{\beta}$ are decreasing to 0 (simulated annealing scheme).

Main theorem. *Let* $F \in B(\Gamma_d(V))$ *and an initial measure* $\nu \in \mathbb{B}(\lambda)$. *Then under the condition*

$$\delta\, e^{\beta b} < const \tag{24}$$

with $b = \sup_{\gamma \in \Gamma_d(V)} \sup_{x \in \gamma} E(x, \gamma \backslash x)$ *we have*

$$\lim_{\beta \to \infty,\ t \to \infty,\ \delta \to 0} \langle F\rangle_{S_{\beta,\delta}([\frac{t}{\delta}])\nu} = \langle F\rangle_{\mu_\infty}, \tag{25}$$

where the measure μ_∞ is defined in Theorem 1, and $\langle F\rangle_{S_{\beta,\delta}([\frac{t}{\delta}])\nu} = \langle S_{\beta,\delta}([\frac{t}{\delta}])\nu, F\rangle$.

Proof. We can write the following

$$\langle F\rangle_{S_{\beta,\delta}([\frac{t}{\delta}])\nu} \;-\; \langle F\rangle_{\mu_\infty} \;=\; (p_\nu, T_{\beta,\delta}([\frac{t}{\delta}])F)_{\mu_\beta} \;-\; (p_\nu, T_\beta(t)F)_{\mu_\beta} \;+$$

$$+(T_\beta(t)p_\nu, F)_{\mu_\beta} \;-\; \langle F\rangle_{\mu_\beta} \;+\; \langle F\rangle_{\mu_\beta} \;-\; \langle F\rangle_{\mu_\infty}.$$

Then, using the results of theorem 1 and the limit relations (7), (18) and (23) we get (25). In addition, the relation (24) follows from the approximation technique (see appendix B, equations (43) and (50)).

Remark. The relation (25) determines the limit over three quantities: $\beta \to \infty,\ t \to \infty, \delta \to 0$. In the approximation technique we used the relation (24) between δ and β:

$$\delta \;=\; \phi(\beta)\; e^{-\beta b} \quad \text{with} \quad \phi(\beta) = O(1) \;\text{ as }\; \beta \to \infty.$$

Unfortunately the relation between t and β is still an open problem. If we have the relation $t = \psi(\beta)$ in an explicit form, then (25) can be rewritten as a limit when $t \to \infty$ under two relations $\beta(t) = \psi^{-1}(t)$ and $\delta(t) = \phi(\beta(t))e^{-\beta(t)b}$.

5 Application to object detection from numerical images

In this section, we address to application in image analysis. We first define a general model for detecting a collection of objects from digital images. we then detail the algorithm derived from the proposed discretization scheme of the described continuous dynamics and finally show some results for trees and flamingos detection.

5.1 The energy function

Let us consider a numerical image on the lattice $I \subset Z^2$ as data. It is defined as follows:

$$\begin{aligned} Y : I \quad &\to \quad \Lambda \subset N \\ s \quad &\mapsto \quad y_s \end{aligned} \tag{26}$$

Each y_s refers to the gray level at pixel s, on the lattice $I = \{1, \cdots, NL\} \times \{1, \cdots, NC\}$, NL (resp. NC) being the number of lines (resp. columns) of the analyzed image. The lattice I is interpreted as a discretization of the

domain V on which the disks configuration is defined. Therefore, we consider configurations of centers of disks $\gamma = \{x_i\} \in \Gamma_d(V)$, where $V = [1/2, NL+1/2] \times [1/2, NC+1/2]$. Each disk in the final configuration represents an object in the image. The hard core distance ϵ is naturally taken to be equal to the data resolution: $\epsilon = 1$ pixel. To define the energy H, we first consider some prior knowledge. We want to minimize the overlap between objects as one object in the image should not be represented by several disks in the estimated configuration.. However, to obtain a more flexible model w.r.t. the data and the object shape, we do not forbid but only penalize overlapping objects. We define a pairwise interaction as follows:

$$\forall \{x_i, x_j\} \in \gamma \times \gamma, H_2(x_i, x_j) = \max\left(0, 1 - \frac{||x_j - x_i||}{2r}\right) \tag{27}$$

where $||.||$ is the Euclidean norm and r is the radius of the underlying disk.

A first order term is then added to the energy function for each object to fit the disk configuration onto the data. It is a data driven term defining the interaction of a configuration of disks with the studied image. We consider that there is an object, modeled by a disk centered at pixel s, in the image, if the gray level values of the pixels inside the projection of the disk onto the lattice are statistically different from those of the pixels in the neighborhood of the disk. To quantify this difference we compute a distance between the associated distributions. Denote $D_1(s)$, the projection of the disk with radius r centered at s onto the lattice, and $D_2(s)$ the surrounding crown:

$$D_1(s) = \{t \in I : ||t - s|| \leq r\} \text{ and } D_2(s) = \{t \in I : r < ||t - s|| \leq r + 1\}. \tag{28}$$

We consider the mean and the variance of the data of these two subsets:

$$\begin{aligned} \mu_1(s) &= \frac{\sum_{t \in D_1(s)} y_t}{\sum_{t \in D_1(s)} 1} \quad \text{and} \quad \mu_2(s) = \frac{\sum_{t \in D_2(s)} y_t}{\sum_{t \in D_2(s)} 1} \\ \sigma_1^2(s) &= \frac{\sum_{t \in D_1(s)} y_t^2}{\sum_{t \in D_1(s)} 1} - \mu_1(s)^2 \quad \text{and} \quad \sigma_2^2(s) = \frac{\sum_{t \in D_2(s)} y_t^2}{\sum_{t \in D_2(s)} 1} - \mu_2(s)^2 \end{aligned} \tag{29}$$

We consider a distance between the distributions in $D_1(s)$ and in $D_2(s)$ which is written as follows:

$$B(s) = \frac{1}{4}\left(\mu_1(s) - \mu_2(s)\right)^2 \sqrt{\sigma_1^2(s) + \sigma_2^2(s)} - \frac{1}{2}\log\frac{2\sigma_1(s)\sigma_2(s)}{\sigma_1^2(s) + \sigma_2^2(s)}. \tag{30}$$

From this distance between the two distributions, a first order energy term is built:

$$\forall x_i \in \gamma,\ H_1(x_i) = \begin{cases} \left(1 - \frac{B(i)}{T}\right) & \text{if} \quad B(i) < T \\ \left(\exp - \frac{B(i) - T}{3B(i)} - 1\right) & \text{if} \quad B(i) \geq T \end{cases} \tag{31}$$

where i is the closest point to x_i on the lattice, and T is a threshold parameter. Finally, under the hard core constraint, the global energy is written as follows:

$$H(\gamma) = \alpha \sum_{x_i \in \gamma} H_1(x_i) + \sum_{\{x_i, x_j\} \in \gamma \times \gamma, i \neq j} H_2(x_i, x_j) \tag{32}$$

where α is a weighting parameter between the data term and the prior.

5.2 Algorithm

The algorithm simulating the process is defined as follows:

- **Computation of the first order term:** For each site $s \in I$ compute $H_1(s)$ from the data
- **Computation of the birth map:** To speed up the process, we consider a non homogeneous birth rate to favor birth where the first order term is low (i.e. where the data tend to define an object):

$$\forall s \in I, b(s) = 1 + 9 \frac{\max_{t \in I} H_1(t) - H_1(s)}{\max_{t \in I} H_1(t) - \min_{t \in I} H_1(t)}. \tag{33}$$

  The normalized birth rate is then given by:

$$\forall s \in I, B(s) = \frac{z b(s)}{\sum_{t \in I} b(s)} \tag{34}$$

  This non homogeneous birth rate refers to a non homogeneous reference Poisson measure. It has no impact on the convergence to the global minima of the energy function but do have an impact on the speed of convergence in practice by favoring birth in relevant locations.

- **Main program:** initialize the inverse temperature parameter $\beta = \beta_0$ and the discretization step $\delta = \delta_0$ and alternate birth and death steps:
  - **Birth step:** for each $s \in I$, if $x_s = 0$ (no point in s) add a point in s ($x_s = 1$) with probability $\delta B(s)$ (note that the hard core constraint with $\epsilon = 1$ pixel is satisfied).
  - **Death step:** consider the configuration of points $x = \{s \in I : x_s = 1\}$ and sort it from the highest to the lowest value of $H_1(s)$. For each point taken in this order, compute the death rate as follows:

$$d_x(s) = \frac{\delta a_x(s)}{1 + \delta a_x(s)}, \tag{35}$$

where:

$$a_x(s) = \exp -\beta \left(H(x/\{x_s\}) - H(x) \right). \tag{36}$$

and kill s ($x_s = 0$) with probability $d(s)$.

– **Convergence test:** if the process has not converged, increase the inverse temperature β and decrease the discretization step δ by a geometric scheme and go back to the birth step. The convergence is obtained when all the objects added during the birth step, and only these ones, have been killed during the death step. Note that we do not respect the condition 24, for saving computation time, but for the applications we address in this paper this approximation does not affect the results.

5.3 Results

The first application we address concerns tree crown extraction from aerial images. The goal of this detection is to count the trees for monitoring the resources and evaluating the biomass. We consider 50cm resolution images of poplars. Other approaches to perform tree detection include template matching [14] or mathematical morphology [15]. Recently, a point process modeling have been proposed, based on an Reversible Jump Markov chain Monte Carlo (RJMCMC) sampler [6]. This model has shown good performances compared to the two previous techniques [16]. The proposed model is very similar to this point process modeling. The main difference lies in the optimization as we consider a birth and death process instead of an RJMCMC scheme. We therefore obtain similar results but we have approximately divided by ten the computation time. Some examples of the obtained results are given on figures 1 and 2. The results are satisfactory. One can remark a few false alarms on figure 1, on the border on the plantation, due to shadows and a few misdetection on figure 2 on small trees for which the chosen radius (3 pixels) is too big.

The second application concerns the counting of flamingo population in order to evaluate its size for studying its dynamic. An extract of the obtained result is given on figure 3 for the initial image and on figure 4 for the detected birds. Almost all the birds have been correctly detected. The full image contains 6128×3920 pixels and has been analyzed in ten minutes on a bi-processor 2GHz PC. This represents a main advantage with respect to more standard optimization techniques based on a RJMCMC sampler [6, 16]. Indeed, the speed of convergence and the computational efficiency of the proposed algorithm allow us to deal with with huge sets of

data in a reasonable time. We have validated the results by comparing them with a detection manually made by an expert from the biological station "La Tour du Valat". On a small part of the image given on figure 3, the expert has detected 128 flamingos. Our result given on figure 5 left, exhibits a good detection rate of $99,2\%$ and 0 false alarm. On a more difficult case, with a low contrast and a perspective effect, we obtain $91,7\%$ of detection rate with $10,9\%$ of false alarms (see figure 5 right), which is quite satisfactory in this context. Some comparative study showing that, as for the tree detection application, our approach overcomes standard techniques such as mathematical morphology or template matching can be found in [18].

6 Conclusion

In this paper, we have proposed a new approach for detecting objects in an image. This approach is based on a birth and death process. We have proved the convergence of the continuous process. We then have described a discretization scheme and proved its convergence to the continuous process. From this general framework, we have proposed a disk model which permits the detection of objects in a given image. Two applications, concerning tree and bird detection, have shown the relevance of the proposed approach. The two main advantages of this technique are its generality and its computational efficiency. The non overlapping prior and the specification of the objects allows us to solve some ambiguities when the objects are very close to eachother. This algorithm has been used by the scientists of La Tour du Valat for evaluating flamingo populations and is evaluated by the National French Inventory for counting trees in an operational context.

Next steps will concern the generalization of the model to a broader class of objects. Taking into account other kinds of objects such as ellipses or rectangles is straightforward. However, it will be interesting to embed some randomness in the definition of objects. Dealing with random radius or more generally random marks associated with the points in the configuration will increase the application domain of this promising approach. To tackle this new generation of models, we are currently working on new dynamics for addressing geometric changes in the configuration such as object dilation, translation, rotation or object splitting and merging.

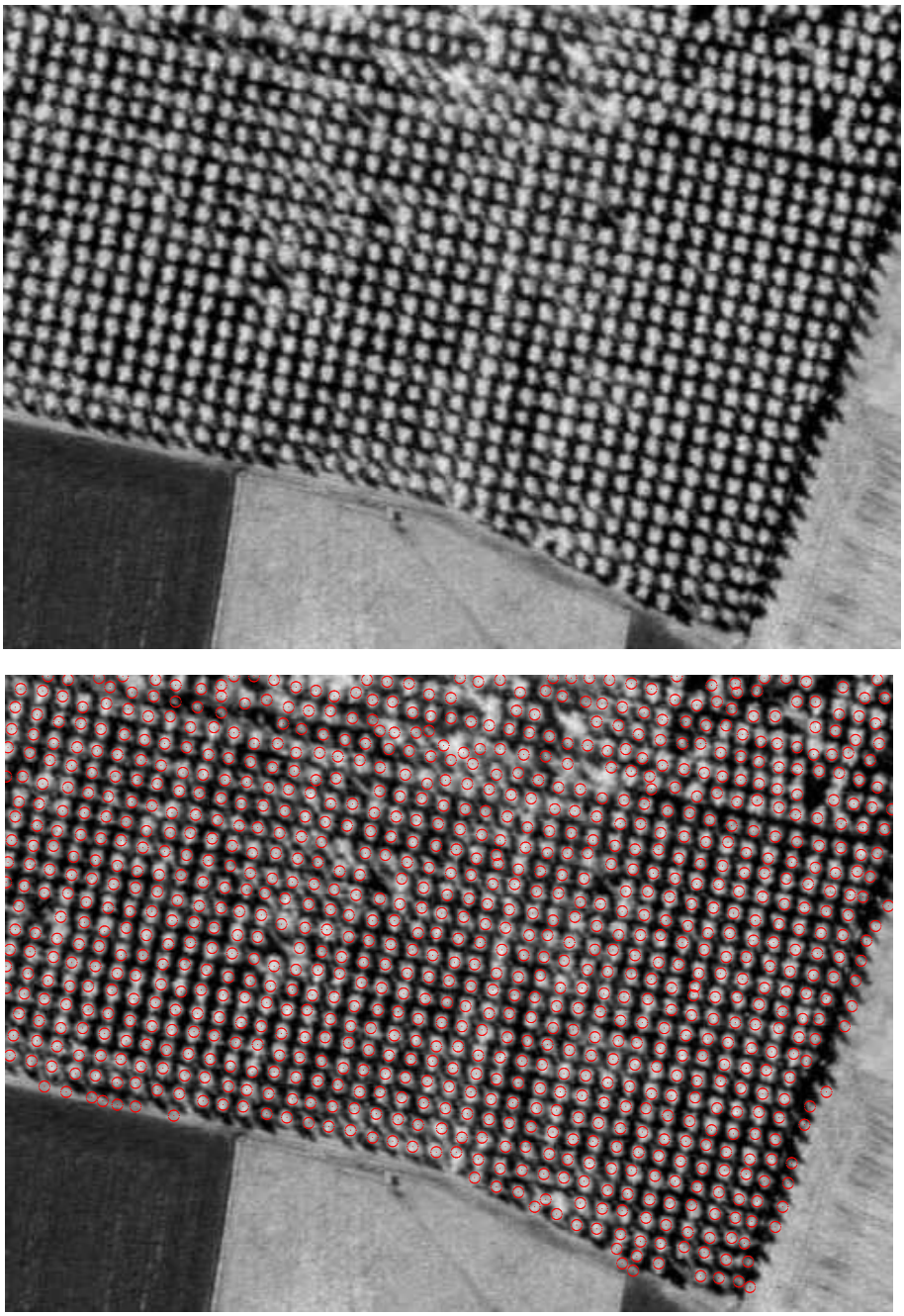

Figure 1: Result on a poplar plantation (top: initial image © IFN, bottom: detected trees)

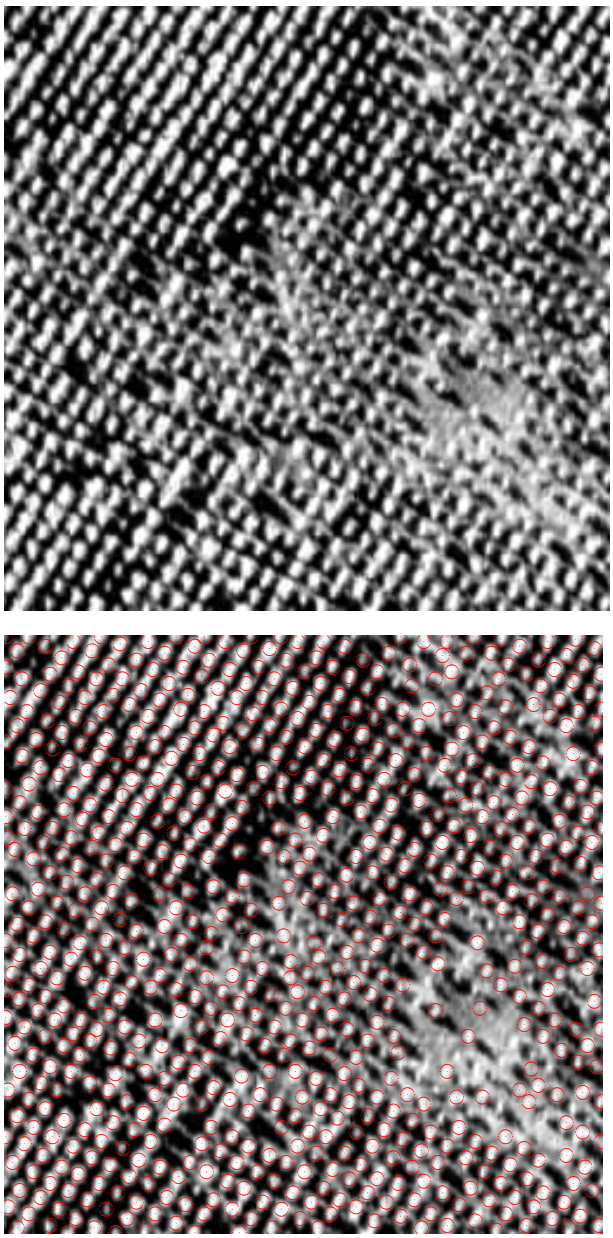

Figure 2: Result on a poplar plantation (top: initial image © IFN, bottom: detected trees)

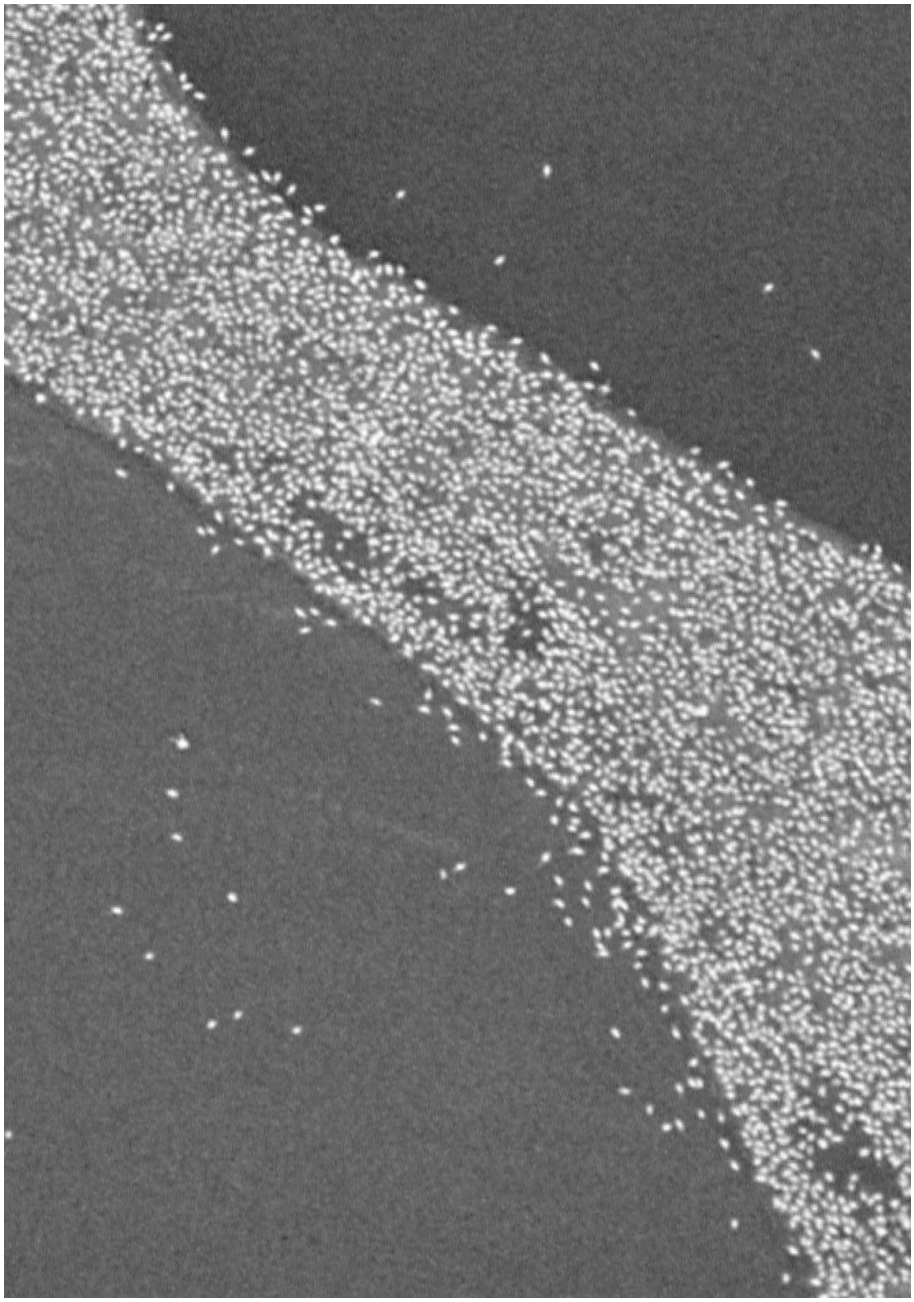

Figure 3: Bird population from France © Station Biologique Tour du Valat

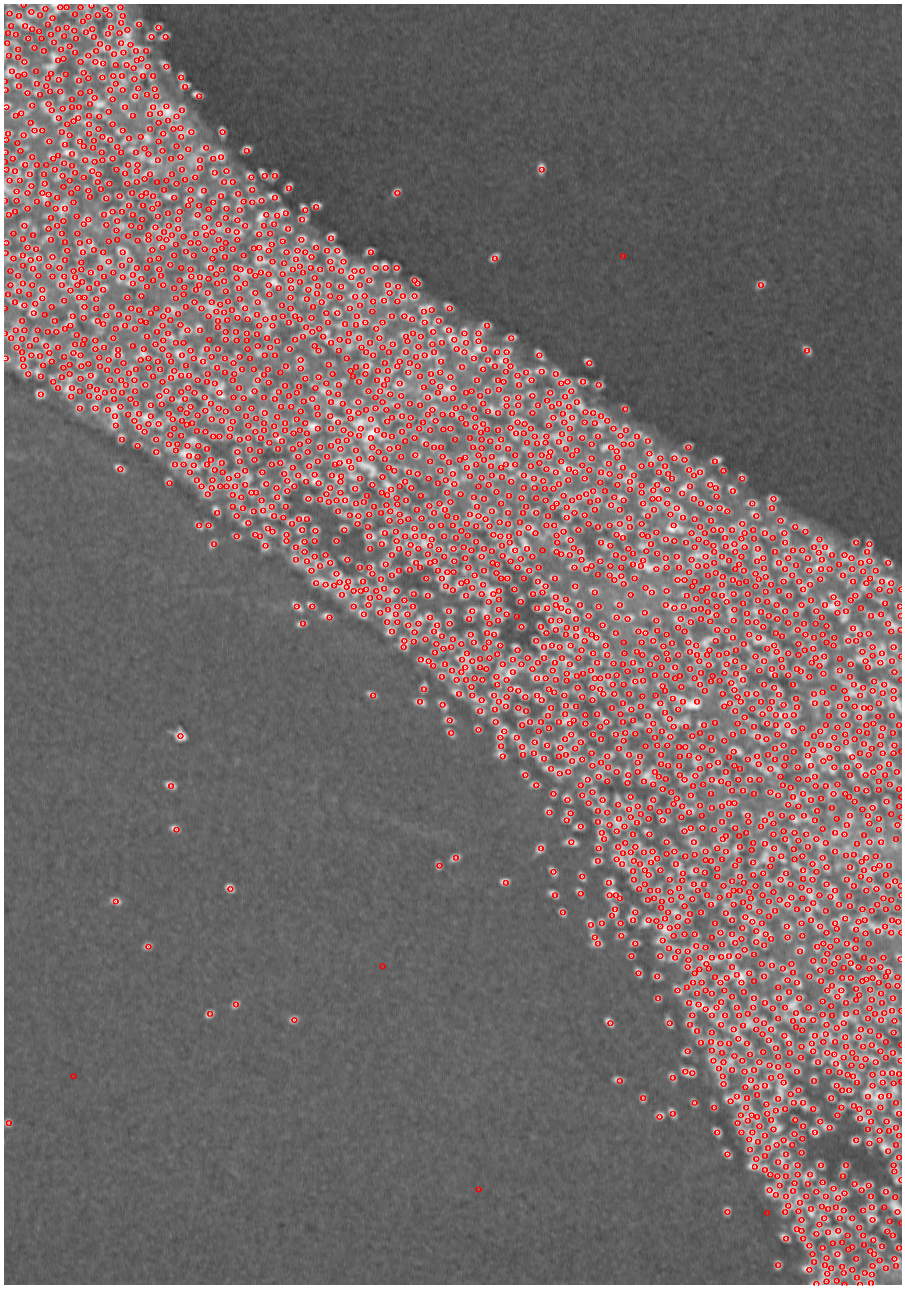

Figure 4: Detected birds from the image shown on figure 3

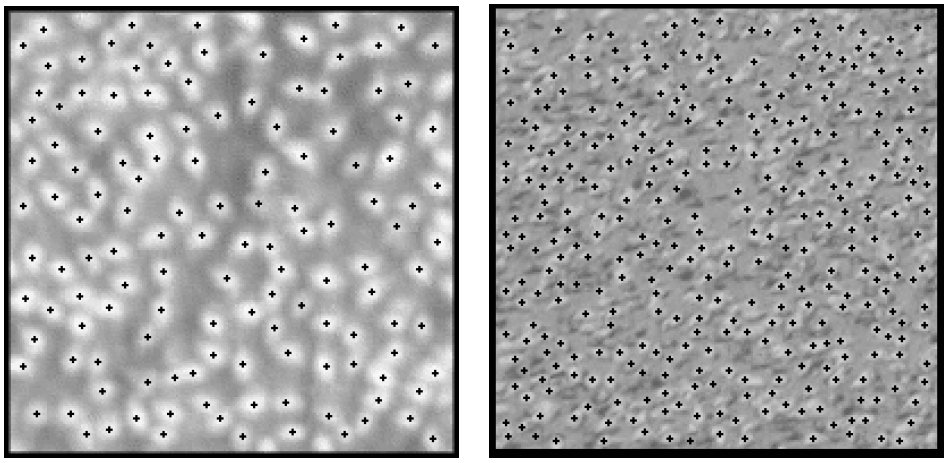

Figure 5: Detected birds on a small part of the image shown on figure 3 (left) and on a more complex image (right) from Mauritania

Appendix A. Proof of theorem 2.

1) The boundedness of L_β is obvious, and then we can define the semi-group $T_\beta(t)$ as the exponent of the operator L_β using the usual expansion for the exponent. Self-adjointness follows from the detailed balance condition and the boundedness of the operator L_β.

2) Since $L_\beta 1 = 0$ we get $e^{tL_\beta}1 = 1$. The second condition in (17) follows from the analogous property of the operators P_β^n and the convergence of the approximation process associated with transition operator P_β (see appendix B).

3) Using the relations

$$T_\beta(t) \;=\; e^{tL_\beta} \;=\; \lim_{n\to\infty}\left(E + \frac{t}{n}L_\beta\right)^n \tag{37}$$

and

$$\left(E + \frac{t}{n}L_\beta\right) f(\gamma) \;=\; \left(1 - \frac{t}{n}\left(\sum_{x\in\gamma} e^{\beta E(x,\gamma\setminus x)} + zV(\gamma)\right)\right) f(\gamma) \;+$$

$$\frac{t}{n}\sum_{x\in\gamma} e^{\beta E(x,\gamma\setminus x)} f(\gamma\setminus x) \;+\; \frac{t}{n}\, z \int\limits_{V(\gamma)} f(\gamma\cup y)\, dy \tag{38}$$

we obtain that, for any $t > 0$ and for large enough n, all coefficients in the decomposition given by equation (38) are positive, and moreover,

$$\left|\left(E + \frac{t}{n}L_\beta\right) f(\gamma)\right| \;\le\; \left|\left(E + \frac{t}{n}L_\beta\right) |f|(\gamma)\right| \;\le\; \sup_\gamma |f(\gamma)|.$$

Thus,

$$\sup_\gamma \left|\left(E + \frac{t}{n}L_\beta\right) f(\gamma)\right| \;\le\; \sup_\gamma |f(\gamma)|, \tag{39}$$

and applying inequality (39) n times we still keep the same bound:

$$\sup_\gamma \left|\left(E + \frac{t}{n}L_\beta\right)^n f(\gamma)\right| \;\le\; \sup_\gamma |f(\gamma)| \quad \text{for all large enough} \quad n \in N.$$

Consequently, $T_\beta(t)$ is a contraction semi-group in $B(\Gamma_d(V))$, i.e.

$$\sup_\gamma |(T_\beta(t)\; f)(\gamma)| \;\le\; \sup_\gamma |f(\gamma)|.$$

Appendix B. Convergence of the approximation process. Proof of theorem 4.

To prove the convergence of the corresponding semi-group

$$T_{\beta,\delta}(t) = P_{\beta,\delta}^{[\frac{t}{\delta}]} \ \rightarrow \ T_\beta(t), \quad \delta \rightarrow 0$$

uniformly on bounded intervals of time t we use the following approximation theorem:

Approximation theorem [17]. *For $n = 1, 2, \ldots$ let T_n be a linear contraction on a Banach space $\mathfrak{L}$, and let δ_n be positive numbers. We set:*

$$L_n \ = \ \frac{1}{\delta_n} \ (T_n - E).$$

Assume that $\lim_{n\rightarrow\infty} \delta_n \ = \ 0$.

Let $\{T(t)\}$ be a strongly continuous contraction semi-group on the Banach space $\mathfrak{L}$ with generator L, and let C be a core for L. Then the following propositions are equivalent:

a) for each $f \in \mathfrak{L}$

$$\|T_n(t)f \ - \ T(t)f\|_{\mathfrak{L}} \ \rightarrow \ 0, \quad as \quad \delta_n \rightarrow 0$$

for all $t \geq 0$ uniformly on bounded intervals;

b) for each $f \in C$

$$\|L_n f \ - \ Lf\|_{\mathfrak{L}} \ \rightarrow \ 0, \quad \text{as} \quad \delta_n \rightarrow 0.$$

We denote by $L_{\beta,\delta}$ the generator of the process $T_{\beta,\delta}$ defined by transition probabilities (21) (homogeneous in time):

$$(L_{\beta,\delta}f)(\gamma) \ = \ \frac{1}{\delta_n} \left(\left(P_{\beta,\delta}f\right)(\gamma) \ - \ f(\gamma) \right) \ = \tag{40}$$

$$\frac{1}{\delta_n} \left(\sum_{\gamma_1 \subseteq \gamma} \prod_{x \in \gamma \setminus \gamma_1} (a_x \delta_n) \prod_{x \in \gamma} \frac{1}{1 + a_x \delta_n} \right.$$

$$\left. \Xi_\delta^{-1}(\gamma_1) \sum_{k=0}^{\infty} \int\limits_{V_k(\gamma_1)} \frac{(z\delta_n)^k}{k!} \ f(\gamma_1 \cup y_1 \cup \ldots \cup y_k) \ dy_1 \ldots dy_k \ - \ f(\gamma) \right) \ =$$

$$\frac{1}{\delta_n} \left(\Xi_\delta^{-1}(\gamma) \prod_{x \in \gamma} \frac{1}{1 + a_x \delta_n} \ f(\gamma) \ - \ f(\gamma) \right) \ +$$

$$\frac{1}{\delta_n}\ \Xi_\delta^{-1}(\gamma\backslash x)\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta_n}\ \sum_{x\in\gamma}a_x\delta_n\ f(\gamma\backslash x)\ +$$

$$\frac{1}{\delta_n}\ \Xi_\delta^{-1}(\gamma)\ z\delta_n\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta_n}\int\limits_{V(\gamma)} f(\gamma\cup\tilde{y})\ d\tilde{y}\ +$$

$$\frac{1}{\delta_n}\ \sum_{\tilde{\gamma}\subseteq\gamma}\Xi_\delta^{-1}(\gamma\backslash\tilde{\gamma})\sum_{k:|\tilde{\gamma}|+k\geq 2}\frac{(z\delta_n)^k}{k!}\prod_{y\in\gamma}\frac{1}{1+a_y\delta_n}\ \prod_{x\in\tilde{\gamma}}a_x\delta_n$$

$$\int\limits_{V_k(\gamma\backslash\tilde{\gamma})} f((\gamma\backslash\tilde{\gamma})\cup y_1\cup\ldots\cup y_k)\ dy_1\ldots dy_k.$$

We take here as a core $C = B(\Gamma_d(V))$ a whole set of bounded functions on Γ_V. Finally, consider the following statement:

Statement *Let us denote*

$$\Delta_\delta f\ =\ L_{\beta,\delta}f\ -\ L_\beta f.$$

Then

$$\sup_\gamma\ |\Delta_\delta f(\gamma)|\ \rightarrow\ 0\quad as\quad \delta\rightarrow 0 \tag{41}$$

for each $f(\gamma)\in B(\Gamma_V)$.

Proof of the statement:

We have for any $x\in\gamma$

$$E(x,\gamma\backslash x)\ \leq\ b, \tag{42}$$

so that

$$a_x\ \equiv\ e^{\beta E(x,\gamma\backslash x)}\ \leq\ e^{\beta b} \tag{43}$$

and let

$$a\ =\ a(\beta)\ =\ \sup_{\gamma\in\Gamma_d(V)}\ \sup_{x\in\gamma}\ a_x\ <\ \infty.$$

Lemma 1. *The normalizing factor* $\Xi_\delta^{-1}(\gamma)$ *from equation (21) can be written as*

$$\Xi_\delta^{-1}(\gamma)\ =\ 1\ -\ z\delta|V(\gamma)|\ +\ O\left(z^2\delta^2\right)\quad \text{as}\quad \delta\rightarrow 0. \tag{44}$$

Proof. Since

$$\Xi_\delta(\gamma) \;=\; 1 \;+\; \sum_{m=1}^{\infty} \frac{(z\delta)^m}{m!} \int\limits_{V_m(\gamma)} dy_1 \dots dy_m \;=\; 1 \;+\; z\delta|V(\gamma)| \;+\; \sum_{m=2}^{\infty} \frac{(z\delta)^m}{m!} \mathcal{V}_m$$

with $\mathcal{V}_m < |V(\gamma)|^m < |V|^m,\ m \geq 2$, we can write

$$\Xi_\delta(\gamma) \;=\; 1 \;+\; z\delta|V(\gamma)| \;+\; O(z^2\delta^2). \tag{45}$$

Thus, (45) implies (44). □

Using the Taylor expansions

$$\ln(1+x) \;=\; x - \frac{x^2}{2}\,\xi', \quad \xi' \in (4/9, 4) \quad \text{as} \quad |x| < \frac{1}{2}, \tag{46}$$

and

$$e^x \;=\; 1 + x + \frac{x^2}{2}\,e^{\xi}, \quad \xi \in (0, x), \tag{47}$$

we have for small enough δ:

$$\frac{1}{1+a_x\delta} \;=\; e^{-\ln(1+a_x\delta)} \;=\; e^{-a_x\delta + \frac{\xi_1}{2}\,a_x^2\delta^2},$$

$$\prod_{x\in\gamma} \frac{1}{1+a_x\delta} \;=\; e^{\,-\delta \sum\limits_{x\in\gamma} a_x + \frac{\xi_1}{2}\,\delta^2 \sum\limits_{x\in\gamma} a_x^2}. \tag{48}$$

Then using (44), (48) and relation $\frac{\xi_1}{2}\delta^2 \sum_{x\in\gamma} a_x^2 = O(a^2\delta^2)$ we have for all small enough δ:

$$\Xi_\delta^{-1}(\gamma) \;\prod_{x\in\gamma} \frac{1}{1+a_x\delta} \;-\; \left(1 - \delta \sum_{x\in\gamma} a_x - z\;\delta\;|V(\gamma)|\right) \;= \tag{49}$$

$$\left(1 - z\delta|V(\gamma)| + O\left((z\delta)^2\right)\right)\; e^{-\delta\sum a_x + \delta^2\frac{\xi_1}{2}\sum a_x^2} \;-$$

$$\left(1 \;-\; \delta\sum a_x \;-\; \delta z|V(\gamma)|\right) \;=\; O\left(\delta^2(a^2+z^2)\right).$$

We assume here that

$$\delta\; a \;=\; \delta\; a(\beta) \;<\; const, \tag{50}$$

which is of course true for any fixed β and small enough δ. Let us write now the expression for $\Delta_\delta f(\gamma)$ using (40) and (13):

$$(\Delta_\delta f)(\gamma) \;=\; \frac{1}{\delta}\left((P_\delta f)(\gamma) \;-\; f(\gamma)\right) \;- \tag{51}$$

$$\sum_{x\in\gamma} a_x(\gamma)\ (f(\gamma\backslash x)-f(\gamma))\ -\ z\int\limits_{V(\gamma)}(f(\gamma\cup y)-f(\gamma))\ dy\ =$$

$$\frac{1}{\delta}\left(\Xi_\delta^{-1}(\gamma)\ \prod_{x\in\gamma}\frac{1}{1+a_x\delta}\ f(\gamma)\ -\ f(\gamma)\right)\ +$$

$$+\ \sum_{x\in\gamma} a_x(\gamma)\ f(\gamma)\ +\ z\ |V(\gamma)|\ f(\gamma)\ +$$

$$\frac{1}{\delta}\ \sum_{x\in\gamma} a_x(\gamma)\delta\ \Xi_\delta^{-1}(\gamma\backslash x)\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta}\ f(\gamma\backslash x)\ -\ \sum_{x\in\gamma} a_x(\gamma) f(\gamma\backslash x)\ +$$

$$\frac{1}{\delta}\ \Xi_\delta^{-1}(\gamma)\ z\ \delta\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta}\int\limits_{V(\gamma)} f(\gamma\cup y)\ dy\ -\ z\int\limits_{V(\gamma)} f(\gamma\cup y)\ dy\ +$$

$$\frac{1}{\delta}\ \sum_{\tilde\gamma\subseteq\gamma}\ \Xi_\delta^{-1}(\gamma\backslash\tilde\gamma)\ \sum_{k:|\tilde\gamma|+k\geq 2}\ \frac{(z\delta)^k}{k!}\ \prod_{x\in\tilde\gamma} a_x\delta\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta}$$

$$\int\limits_{V_k(\gamma\backslash\tilde\gamma)} f(\gamma\backslash\tilde\gamma\cup y_1\cup\ldots\cup y_k)\ dy_1\ldots dy_k.$$

Estimation of the coefficients in (51):

Let us estimate the terms with $f(\gamma),\ f(\gamma\backslash x),\ f(\gamma\cup y)$ in (51) separately.

Using (49) and boundedness of $f(\gamma)$ we have:

$$\left|\frac{1}{\delta}\left(\Xi_\delta^{-1}(\gamma)\ \prod_{x\in\gamma}\frac{1}{1+a_x\delta}\ f(\gamma)\ -\ f(\gamma)\right)\ +\right.$$

$$\left.\left(\sum_{x\in\gamma} a_x(\gamma)\ +\ z|V(\gamma)|\right) f(\gamma)\right|\ =\ O(\delta(a^2+z^2)),\ \delta\to 0. \tag{52}$$

Analogously, we can estimate the coefficients before $f(\gamma\backslash x)$ and $f(\gamma\cup y)$:

$$\left|\sum_{x\in\gamma} a_x\left(\Xi_\delta^{-1}(\gamma\backslash x)\ \prod_{y\in\gamma}\frac{1}{1+a_y\delta}\ -\ 1\right)\ f(\gamma\backslash x)\right|\ =$$

$$\left|\sum_{x\in\gamma} a_x\left(\left(1-z\delta|V(\gamma\backslash x)|+O(z^2\delta^2)\right)\ e^{-\delta\sum a_x+\frac{\xi_1}{2}\delta^2\sum a_x^2}\ -\ 1\right)\ f(\gamma\backslash x)\right|\ \leq$$

$$a|\gamma|\ \left(\delta(a|\gamma|+z|V|)\ +\ \delta^2\ (A_2|\gamma|+A_3z^2|V|^2)\right)\ K_f\ =\ O(\delta a(z+a)). \tag{53}$$

And

$$z\left|\int\limits_{V(\gamma)}\left(\Xi_\delta^{-1}(\gamma)\ \prod_{x\in\gamma}\frac{1}{1+a_x\delta}\ -\ 1\right) f(\gamma\cup y)\ dy\right|\ \leq$$

$$z|V|\left(\delta(a|\gamma|+z|V|)\ +\ \delta^2(A_2|\gamma|+A_3z^2|V|^2)\right)\ K_f\ =\ O(\delta z(z+a)). \tag{54}$$

Let us estimate now the last term in (51). Using that

$$\Xi_\delta^{-1}(\gamma)\ \le\ 1,\quad \prod_{y\in\gamma}\frac{1}{1+a_y\delta}\ \le 1,$$

and $\sup_\gamma|f(\gamma)|<K_f$ we have

$$\frac{1}{\delta}\sum_{\tilde\gamma\subseteq\gamma}\Xi_\delta^{-1}(\gamma\backslash\tilde\gamma)\sum_{k:|\tilde\gamma|+k\ge2}\frac{(z\delta)^k}{k!}\prod_{x\in\tilde\gamma}a_x\delta\prod_{y\in\gamma}\frac{1}{1+a_y\delta}$$

$$\int\limits_{V_k(\gamma\backslash\tilde\gamma)}f(\gamma\backslash\tilde\gamma\cup y_1\cup\ldots\cup y_k)\ dy_1\ldots dy_k\ \le$$

$$\frac{K_f}{\delta}\sum_{\tilde\gamma\subseteq\gamma}(a\delta)^{|\tilde\gamma|}\sum_{k\ge0:|\tilde\gamma|+k\ge2}\frac{(z|V|\delta)^k}{k!}\ \le$$

$$\frac{K_f}{\delta}\ a^N\sum_{\substack{m=0,\ldots,|\gamma|\\ k\ge0:m+k\ge2}}C^m_{|\gamma|}\ \delta^m\ \frac{(z|V|\delta)^k}{k!}\ =$$

$$\frac{K_f}{\delta}\ a^N\ \left(e^{z|V|\delta}(1+\delta)^{|\gamma|}-1-|\gamma|\delta-z|V|\delta\right)\ =\ O(\delta+\delta z+\delta z^2). \tag{55}$$

Here we used that

$$(1+\delta)^{|\gamma|}\ =\ e^{|\gamma|\ln(1+\delta)}\ =\ e^{\delta|\gamma|}\ +\ O(\delta^2)\quad \text{for small } \delta>0.$$

Finally, from (51), (52), (53), (54), (55) it follows that for any $f(\gamma)\in B(\Gamma_d(V))$

$$\sup_\gamma|\Delta_\delta f(\gamma)|\ \to\ 0\quad\text{as}\quad\delta\to0,$$

and consequently,

$$\|L_{\beta,\delta}f\ -\ L_\beta f\|_{B(\Gamma_d(V))}\ \to\ 0\quad\text{as}\quad\delta\to0.$$

The statement is proved and relation (41) immediately implies convergence (22) of the semi-groups in the uniform norm of the space $\mathcal{L}$ by the above approximation theorem. Theorem 4 is proved.

References

[1] L. Bertini, N. Cancrini, and F. Cesi. The spectral gap for a Glauber-type dynamics in a continuous gas. *Ann. Inst. H. Poincare Probab. Statist.*, vol. 38, pp. 91–108, 2002.

[2] Y.G. Kondratiev, R.A. Minlos, and E.A. Zhizhina. One particle subspace of the Glauber dynamics generator for continuous particle systems. *Reviews in Math. Physics*, vol. 16, no. 9, pp. 1073–1114, 2004.

[3] R. Stoica, X. Descombes, and J. Zerubia. A Gibbs point process for road extraction in remotely sensed images. *International Journal of Computer Vision*, vol. 57, no. 2, pp. 121–136, 2004.

[4] C. Lacoste, X. Descombes, and J. Zerubia. Point processes for unsupervised line network extraction in remote sensing. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1568–1579, October 2005.

[5] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, vol. 72, no. 2, pp. 107–132, 2007.

[6] G. Perrin, X. Descombes, and J. Zerubia. 2D and 3D vegetation resource parameters assessment using marked point processes. in *Proc. International Conference on Pattern Recognition (ICPR)*, Hong-Kong, August 2006.

[7] P.J. Green. Reversible jump Markov chain Monte-Carlo computation and Bayesian model determination. *Biometrika*, vol. 57, pp. 97–109, 1995.

[8] M.V. Fedoriuk. Saddle point method. *Chapter II.4 : The Laplace method for multiple integrals* Moscow Nauka, 1977, in Russian.

[9] A. Erdelyi. Asymptotic expansions. Dover, 1956

[10] B.D. Ripley. Modeling spatial patterns. *J. R. Statist. Soc. B*, vol. 39, pp. 172–212, 1977.

[11] T.M. Liggett. Interacting Particle Systems. Springer-Verlag New York Inc., 1985.

[12] M. Reed and B. Simon. Methods of modern mathematical physics, Vol. 4. Academic Press NY, 1978.

[13] C. Preston. Spatial birth-and-death processes. *Bull. Int. Statist. Inst.*, vol. 46, pp. 377–391, 1977.

[14] R. Pollock. The Automatic Recognition of Individual Trees in Aerial Images of Forests based on a Synthetic Tree Crown Image Model. *Ph D thesis, University of British Colombia*, Vancouver, Canada, 1996.

[15] F. Gougeon. A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian Journal of Remote Sensing*, 21(3), pp. 274-284, 1995.

[16] M. Eriksson, G. Perrin, X. Descombes, and J. Zerubia. A comparative study of three methods for identifying individual tree crowns in aerial images covering different types of forests. in *Proc. International Society for Photogrammetry and Remote Sensing (ISPRS, commission I symposium)*, Marne La Vallee, France, July 2006.

[17] S.N. Ethier and T.G. Kurtz. Markov processes: characterization and convergence. John Wiley and sons, N.-Y., 1986.

[18] S. Descamps and X. Descombes and A. Béchet and J. Zerubia. Détection de flamants roses par processus ponctuels marqués pour l'estimation de la taille des populations. *INRIA Research Report*, No 6328, oct. 2007, (in French).